NATIONAL UNIVERSITY OF MODERN LANGUAGES
ISLAMABAD
DEPARTMENT OF COMPUTER SCIENCE
Open-Handed LAB Exam 2
Artificial Intelligence BSCS 6th Evening
Time: 2: 30 Hours

Your task is to build and evaluate Naive Bayes and K-Nearest Neighbors (KNN) classification models.

Data Understanding (5 Marks)

Load the dataset using Python (Pandas).

Display:

- first 10 rows
- dataset shape
- summary statistics
- Plot a scatter plot of feature1 vs feature2 colored by the class label.

Preprocessing (15 Marks)

- Check for missing values.
- **Normalize or standardize the features (explain why normalization is required for KNN).**
- **Split the dataset into training (70%) and testing (30%).**

Naive Bayes Classification (10 Marks)

- Train a Gaussian Naive Bayes classifier on the training set.
- Predict class labels for the test set.
- Evaluate using:
- Accuracy
- Precision, Recall, F1-score
- **Interpret the results:**

K-Nearest Neighbors (KNN) Classification (10 Marks)

- Train a KNN model using:
- k = 3
- Euclidean distance

- Evaluate using the same metrics as Naive Bayes.
- Compare KNN performance for k = 3, 5, 7
- Discuss:
    - How does increasing k affect performance?
    - Which classifier (NB or KNN) performs better and why?

Comparison & Critical Analysis (5 Marks)

- Write a short technical report (8–10 sentences) comparing:
- Give the classification Report of each classifier
- When Naive Bayes is more suitable
- When KNN is more suitable
- Which model fits the given dataset better
- Strengths & weaknesses of both algorithms in binary classification tasks